Noble Saji Mathews

Education

2023–Present **University of Waterloo, Canada**,
MMath Computer Science, *CGPA – 96/100*,
Focus: AI for Software Engineering.

2019–2023 **Indian Institute of Technology, Tirupati**,
Bachelor of Technology, *CGPA – 9.53/10*,
Thesis: Reinforcement Learning for Combinatorial Optimization.

Awards

ACM SIGSOFT Distinguished Paper Award @ MSR 2024

UWaterloo International Masters Award of Excellence, DRP Mentorship Award

Velocity Awarded Up-Start funding to commercialize research (Fuzz-Testing)

Mitacs Globalink Research Internship & Graduate Fellowship Award

IIT-T Governor's Medal for All-Round performance in Undergrad

MLH Software Engineering Fellowship @ Solana (Web3)

GitHub Externship @ DeepSource Labs (Source Code Analysis)

Talks

2nd Nov 2023 **LLMs In Action: Towards Automating Bug Fixes And Enhancing Code Security**, *Research Discovery Days: Mitigating Risks of AI*, Research Spotlight.

Publications

ASE'24 **Test-Driven Development and LLM-based Code Generation**.
- **Noble Saji Mathews**, Meiyappan Nagappan
- To appear in Proceedings of the 39th IEEE/ACM International Conference on Automated Software Engineering

ICSE'24 **FuzzSlice: Pruning false positives in static analysis warnings through function-level fuzzing**.
- Muralli Anirudhan, **Noble Saji Mathews**, Meiyappan Nagappan
- Proceedings of the 46th IEEE/ACM International Conference on Software Engineering

MSR'24 **Whodunit: Classifying Code as Human Authored or GPT-4 generated-A case study on CodeChef problems**.
- Joy Idialu, **Noble Saji Mathews**, Rungroj Maipradit, Joanne Atlee, Meiyappan Nagappan
- Proceedings of the 21st International Conference on Mining Software Repositories. 2024

JSS **On the impact of multiple source code representations on software engineering tasks—An empirical study**.
- Karthik Swarna, **Noble Saji Mathews**, Dheeraj Vagavolu, Sridhar Chimalakonda
- Journal of Systems and Software 210 (2024): 111941

ASE'23 **COMEX: A Tool for Generating Customized Source Code Representations**.
- Debeshee Das*, **Noble Saji Mathews***, Alex Mathai, Srikanth Tamilselvam, Kranthi Sedamaki, Sridhar Chimalakonda, Atul Kumar
- In 2023 38th IEEE/ACM International Conference on Automated Software Engineering (ASE) (pp. 2054-2057). IEEE

EASE'22 **Exploring Security Vulnerabilities in Competitive Programming**.
- Debeshee Das*, **Noble Saji Mathews***, Sridhar Chimalakonda
- Proc. International Conference on Evaluation and Assessment in Software Engineering (EASE), June 2022

IEEE VIS'21 **AiR: An Augmented Reality Application for Visualizing Air Pollution**.
- **Noble Saji Mathews**, Sridhar Chimalakonda, and Suresh Jain
- Proc. IEEE Visualization Conference (VIS), October 2021

CSCW'20 **YTCoder - Towards Turning YouTube into a Development Environment**.
- **Noble Saji Mathews**, Sridhar Chimalakonda, and Akhila Sri Manasa Venigalla
- Proc. ACM Computer Supported Cooperative Work and Social Computing (CSCW), October 2020

* **Shared first authorship**

Select Preprints

IBM Research **CodeSAM: Source Code Representation Learning by Infusing Self-Attention with Multi-Code-View Graphs**.
- Alex Mathai*, Kranthi Sedamaki*, Debeshee Das*, **Noble Saji Mathews***, Srikanth Tamilselvam, Sridhar Chimalakonda, Atul Kumar

Undergrad Research **Statically Detecting Buffer Errors in Cross-language Android Applications Written in Java and C/C++**.
- Kishanthan Thangarajah, **Noble Saji Mathews**, Michael Pu, Meiyappan Nagappan, Yousra Aafer, Sridhar Chimalakonda

Mitacs Whitepaper **DEVAA: A Framework for Detecting Exploitable Vulnerabilities in Android Applications**.
- **Noble Saji Mathews**, Meiyappan Nagappan, Sridhar Chimalakonda

Experience

Mar 2024 – **Current** **Mayil AI**, *Ontario*, Canada, (Bug Localization AI Startup).
- **Position: Co-Founder / Lead AI & Research @ Mayil**
  - Designed and built an **enterprise-grade solution** for on-prem helm deployment. The product was designed to handle the challenges of working with **real-world software tickets** and assist developers without disrupting existing workflows
  - Served as the primary research lead and implemented in-house Agentic Systems to leverage LLMs while reducing hallucinations using experience from **Machine Learning for Software Engineering Research**. Managed multiple successful pilots

Jan 2024 – **Current** **Velocity Incubator**, *Ontario*, Canada, (Efficient Fuzz-Testing).
- **Position: Up-Start Mentee / Lead Code Analysis @ FuzzSlice**
  - Worked on commercializing research after validation and feedback from popular Open-Source projects like Tmux & OpenSSH-portable, followed by market analysis
  - Built the **Static Analysis** for Slice Creation and further refined the process to address speed and memory constraints when analyzing large projects like the Linux kernel
  - Currently designing a commit-level **cloud fuzzing** setup to secure more projects

May 2021 – **Current** — **University of Waterloo SWAG Lab**, *Canada*, (Software Engineering Research).

- **Position: Mitacs Globalink Graduate Fellow**
  - Currently working under the supervision of Dr. Mei Nagappan exploring **AI for Software Engineering**. Particularly LLMs for Code Generation and Testing.
  - Built a framework to help reduce false positives given a time budget using **Dynamic Analysis** techniques like directed fuzzing [Docker Image]
- **Position: Mitacs Globalink Undergraduate Research Intern**
  - Worked on **Software Security** and Source Code Analysis. Built a framework for detecting and exploiting common **Android Vulnerabilities** such as Cross Site Scripting, Intent Reflection and Fragment Injection [Framework]
- **Position: Research Collaborator**
  - Built a Java framework along with IDE Plugins, CI/CD to automate the detection of exploitable vulnerabilities across the Java Native Interface through **Cross-Language Static Analysis** [Java Package]

July 2022 – **Mar 2024** — **IBM Research**, *Bangalore*, India, (AI for Code Lab).

- **Position: Research Collaborator**
  - Built a python package to extract custom **code-views** from Java/C# Code. Implemented and designed a custom RDA algorithm required for subsequent research while benchmarking the quality & performance against other tools like SOOT [Repo]
  - Collaborated on the development of CodeSAM to enhance **source-code representation learning** using self-attention masks, trained SLMs outperformed existing baselines in ML for SE Tasks

May 2023 – **Jan 2024** — **Sudoviz**, *California*, US, (Code Security Startup).

- **Position: AI Engineer**
  - Led the research and development of **Agentic AI solutions** to build an AI platform that empowers AppSec teams to detect and remediate code vulnerabilities
  - Utilized LLMs and Airflow executors to build a Text-to-Cypher solution for data lake, Chat with Repo backend and Vulnerability Scanning prototype

Jan 2022 – **May 2023** — **DeepSource**, *Bangalore*, India, (Static Analysis: Code Security & Reliability).

- **Position: Language Engineering Intern**
  - Offered a part-time position with the language engineering team to continue working on **transitive dependency chain** analysis after GitHub fellowship
  - Worked on creating and documenting "Depend" and integrating the dependency analyzer into the development workflow at the organization [Python Package]
- **Position: GitHub Extern**
  - Worked on developing a language-agnostic dependency analyzer "Depend" which is currently an internal scanner at DeepSource

Apr 2020 – **Apr 2023** — **RISHA Lab IIT Tirupati**, *(Research in SE & HCI)*.

- **Position: Student Researcher**
  - Actively involved in undergrad research under the guidance of Dr. Sridhar Chimalakonda spanning several areas **(XR Visualization, EdTech Solutions, HCI targetted at developers)**. During my final years of undergrad, my research targetted Code analysis and ML in SE through Code Graphs & efficient Source Code Representations [Repo]

Oct 2022 – **Dec 2022** — **Solana Labs**, *San Francisco*, California, (Web3 & Decentralized Finance).

- **Position: MLH Software Engineering Fellow**
  - Awarded Major League Hacking fellowship to collaborate on real-world software projects
  - Worked with the Solana Blockchain and on building a Python SDK for Zeta - DeFi derivatives built on Solana

Outreach

May 2024 – **Current**
Women in Math Directed Reading & Research Programs, *UWaterloo*, Canada.
- **Position: Mentor**
  - Served as a mentor for the WiM Directed Reading & Research Programs, guiding underrepresented undergraduates without prior experience in research
  - Taught students about building generative AI applications, walked them through recent research and explored Evaluation of **Agentic Systems in Software Engineering**

Aug 2020 – **Aug 2022**
EPICS Purdue University, *West Lafayette*, Service Learning Program.
- **Position: Local Team Coordinator**
  - Collaborated with the team at Purdue University to create and demonstrate STEM experiments to enhance the scientific temperament of students across different grades
  - This project aimed to design a **mobile science laboratory** that brings experiments, complete with the necessary materials, to rural schools lacking such facilities

Jul 2019 – **Feb 2022**
Idea Square Lab, *Innovation Hub*, IIT Tirupati.
- **Position: Tech Lead**
  - Developed a cross-platform **tutoring platform** to support real-time one-to-many mentor-mentee interactions. Also prototyped an online testing system for students.
  - Collaborated in developing an interactive platform to accompany a DIY **electronics learning kit** supervised by Dr. T S Natarajan, Professor Department of Physics.
  - Founded the student-run campus newsletter "UDAAN" while a member of the Literary Affairs Council and served as an editor and technical lead